Screening for Soluble Expression of Recombinant Proteins Using the RTS 100 HY in a 96-Well Format

Didier Busso[1], Rosalind Kim[2], and Sung-Hou Kim[1, 2]*
[1]Department of Chemistry, University of California, Berkeley, California, USA
[2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA
**Corresponding author: djbusso@lbl.gov*

Didier Busso

For structural and functional genomics programs, new high-throughput methods to obtain well-expressing and highly soluble proteins are essential. Here, we describe a rapid procedure to express recombinant proteins in the Rapid Translation System RTS 100, an *Escherichia coli* cell-free system using a 96-well format. The identification of the soluble proteins is performed by the dot-blot procedure using an anti-histidine tag antibody.

Introduction

During the last decade, many genomes from both prokaryotes and eukaryotes have been completely sequenced, and more are under way (www.tigr.org/tdb). Based on sequence analysis, a large fraction of genes have unknown cellular and / or molecular functions. One major challenge is to assign biological function and to elucidate the mechanism of action of those open reading frames. The three-dimensional structure of a protein can often provide functional clues, primarily by detecting structural homology with a protein of known function even when sequence homology is low [1]. To face this challenge, several centers worldwide are undertaking structural genomics initiatives ([2] and www.rcsb.org/pdb/strucgen.html).

For structural studies, high yields of soluble recombinant proteins are required. Due to fast growth, easy handling and low cost, *Escherichia coli* is the principal expression system of choice [3, 4].

Nevertheless, recombinant proteins produced in *E. coli* often accumulate as insoluble aggregates. Changing parameters such as temperature, additives, induction conditions, or the addition of fusion partners may alter the behavior of recombinant proteins. Therefore, the development of fast, effective screening methods for expression and solubility is necessary. A faster and more convenient approach to screen the expression of recombinant proteins compared to classical *in vivo* system is the Rapid Translation System RTS 100 [5].

Here, we present the results obtained from screening 24 open reading frames of unknown function from different microorganisms. In order to screen different variables that may interfere with solubility, we expressed the recombinant proteins with a 6-histidine tag, either N-terminal or C-terminal (Nt His and Ct His, respectively) using the RTS 100 at two temperatures (25 °C and 30 °C). The screening was done in a 96-well format. The expression and protein solubility was characterized performing dot blots using an antibody directed against the histidine tag.

We designed a rapid method that allows

- the characterization of soluble candidates from a large number of genes or from a large number of variants,
- complete automation of the process and
- an immediate scale-up of protein expression for the selected candidates.

Materials and Methods

Expression-vector construction

The genes of interest (Table 1) were amplified by polymerase chain reaction (PCR) from the corresponding genomic DNA using a forward primer introducing a *Nde* I site and a reverse primer introducing a *Xho* I site, a stop codon, and a *Bam* HI site. After quantification, 1 µg of each PCR product was digested using either *Nde* I and *Xho* I, or *Nde* I and *Bam* HI, and then purified using a DNA gel extraction kit. To express the recombinant proteins under control of the T7 promoter with either a N-terminal tag (Nt His) or a C-terminal tag (Ct His), the corresponding digested PCR products were subcloned in frame into pIVEX2.4b-Nde and pIVEX2.3-MCS (Roche Applied Science), respectively. All DNA preparations of the recombinant expression vectors were made using either glass-fiber spin tubes or spin miniprep kits.

Table 1: List of proteins selected for *in vitro* expression-solubility screening. The protein name is based on the two letters corresponding to the organism's DNA (amplified by PCR), and the three digits correspond to the last three digits of the *Mycoplasma pneumoniae* GenInfo Identifier (GI) number. Two proteins with the same three digits and different letters represent two homologues of the corresponding *M. pneumoniae* open reading frame from different organisms.

Protein	Organism	GI Number	Molecular Weight (kDa)
AA967	*Aquifex aeolicus*	2983573	36.2
BS217	*Bacillus subtilis*	2633884	16.5
BS676		2635300	19.1
BS742		2632368	20.1
BS967		2633961	35.7
BS994		2633216	18.6
EC731	*Escherichia coli*	1790387	11.7
EC836		1787757	15.1
MG217	*Mycoplasma genitalium*	3844824	17.8
MG222		3844817	44.6
MP003	*Mycoplasma pneumoniae*	1674003	13.6
MP004		1674004	16.8
MP034		1674034	16.2
MP182		1674182	13.4
MP217		1674217	16.3
MP235		1674235	17.7
MP278		1674278	29.0
MP865		1673865	31.0
MP883		1673883	15.4
MP920		1673920	40.5
MP958		1673958	22.4
MP967		1673967	36.6
TM142	*Thermotoga maritima*	4981173	31.2
TM915		4981625	15.8

In vitro protein expression in 96-well plates

In vitro screening for protein expression and solubility was carried out using the cell-free protein synthesis RTS 100 *E. coli* HY Kit (Roche Applied Science). For expression, the reaction was prepared according to Roche Applied Science's instructions and dispensed in 50-µl aliquots into a PCR 96-well reaction plate.

After addition of 150 ng of DNA template, the reactions were incubated for 4 hours either at 25 °C or 30 °C in the Proteomaster Instrument (Roche Applied Science) or a PCR device. After incubation, the soluble proteins were obtained by centrifuging the total reaction mix in the 96-well plate at 3,000 *g* for 20 minutes at 4 °C.

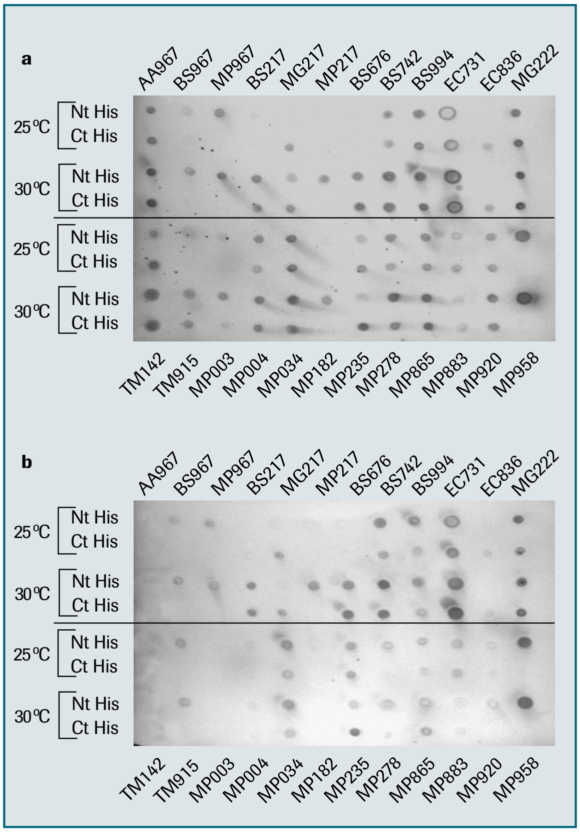
Dot-blot procedure

To test for recombinant protein expression, 1.5 µl of total protein extract and soluble proteins were spotted onto Hybond™ ECL™ nitrocellulose membrane (Amersham Pharmacia Biotech). The prepared membranes were air-dried, then incubated with anti-His monoclonal mouse antibody (Roche Applied Science) diluted 1:5,000 in 3 % powdered milk for 1 hour. Membranes were then washed three times for 10 minutes in TBST (50 mM Tris-HCl pH 7, 250 mM NaCl, 0.05 % Tween 20). The secondary antibody, anti-mouse IgG-horseradish peroxidase conjugate (Roche Applied Science) was applied at 1:2,000 dilution in TBST + 3 % bovine serum albumin (BSA) for 20 minutes. After washing three times for 10 minutes in TBST, blots were developed with 3,3',5,5'-tetramethylbenzidine (TMB) solution.

Results and Applications

Screening for conditions where soluble protein is expressed

Twenty-four ORFs encoding hypothetical proteins ranging from 11.7 kDa to 44.6 kDa from various microorganisms (Table 1) were used to identify conditions allowing soluble expression *in vitro* in a 96-well format. Recombinant proteins carrying either a N-terminal histidine tag or a C-terminal histidine tag were expressed in a 50-µl reaction using the RTS 100 HY Kit at two temperatures (25 °C and 30 °C). Dot blots corresponding to total protein and soluble protein expression are shown



➲ Figure 1:
Dot blots of recombinant proteins expressed *in vitro* at 25 °C and 30 °C. Protein samples were spotted onto nitrocellulose membrane. His_6-tagged proteins were probed using an anti-histidine antibody (see Materials and Methods). a) Total protein from reaction mix b) Soluble protein from the centrifuged reaction mix (Nt His: N-terminal 6 histidine tag, Ct His: C-terminal 6 histidine tag)

in Figures 1a and 1b, respectively. Of the 24 ORFs, three exhibit low expression level (BS967, EC836, and MP883) although these proteins were expressed at a detectable level *in vivo.* Optimizing the *in vitro* expression conditions may increase the expression level of these proteins. Almost 50 % of the expressed proteins show detectable soluble expression *in vitro,* correlating to data obtained *in vivo* (data not shown). Moreover, the location of the histidine tag may affect not only solubility (see Nt His TM915 , MP278, and Ct His MG217), but also expression (see Nt His BS967, MP967, MP217, MP003, MP182, MP958, and Ct His EC836) of recombinant proteins *in vitro.* Lowering the temperature does not seem to increase soluble expression of recombinant proteins *in vitro* as compared to *in vivo* [4].

Using a dot-blot procedure for protein identification via a universal tag allows good qualitative detection of clones expressing soluble protein. The method has not been optimized for the quantitative measurement of protein concentrations. Nevertheless, since the *in vitro* expression reaction was performed in the presence of the same amount of plasmid DNA template, a quick comparison of the expression level may be done by scanning or by visual check.

Since it appears that most of the proteins behave similarly *in vivo* and *in vitro* ([5], [6], Shigeyuki Yokoyama, RIKEN Genomic Sciences Center, personal communication; data not shown), the cell-free system may be used (instead of the time-consuming *in vivo* system) to quickly identify the constructs and parameters that yield soluble expression.

The convenience of the RTS 100 and the speed with which results are obtained are highly compatible with structural genomics expectations. Moreover, the scalability of the RTS allows rapid scale-up protein expression of selected candidates, yielding expression levels compatible with structural studies.

The work described here was supported by the National Institutes of Health GM 62412.

References

1. Zarembiski TI et al. (1999) Proc Natl Acad Sci USA 95: 15189 – 15193
2. Brenner SE (2001) Nat Genet 2: 801 – 809
3. Baneyx F (1999) Curr Opin Biotechnol 10: 411 – 421
4. Hannig G and Makrides SC (1998) Trends Biotechnol 16: 54 – 60
5. Monchois V et al. (2002) Biochemica 1: 22 – 23
6. Kigawa T et al. (2002) J Struct Funct Genomics 2: 29 – 35

Order INFO

Product	Pack Size	Cat. No.
Rapid Translation System RTS 100 *E. coli* HY Kit	24 reactions 96 reactions	3 186 148 3 186 156
Rapid Translation System RTS pIVEX His-tag Vectors	5 vectors	3 253 538
Anti-His6	100 mg	1 922 416
RTS Proteomaster Instrument	1 instrument plus inserts	3 064 859

Imprint

Roche Diagnostics GmbH
Roche Applied Science
Sandhofer Straße 116
D - 68305 Mannheim

Chief Editors: Dr. Friedhelm Huebner, Dr. Ulrich Jellinghaus
Editorial Coordinator: Padma Rao

Springer Verlag GmbH & Co. KG
Science Communications
PD Dr. Beate Fruhstorfer,
Ulrike Hafner, Ursula Hilpert,
Dr. Friederike Holthausen, Sabine Jost,
Dr. Ingeborg König, Dr. Christine Leist,
Sandra Thake (resp.), Teresa Windelen

Biochemica is published quarterly by Roche Applied Science and Springer Verlag and is intended to disseminate information that is of interest and useful to the research community.

We always welcome your comments and suggestions. Contact your local affiliate for a FREE subscription or to change your address.

We welcome manuscripts on any life science topic in which Roche Applied Science's products are used in novel techniques, improvements of common techniques, or simplified protocols. Manuscripts and correspondence concerning editorial matters should be addressed to:

Springer Verlag GmbH & Co. KG
Editorial Office Biochemica
Sandra Thake
Tiergartenstraße 17
D - 69121 Heidelberg
E-mail: thake@springer.de

Biochemica 4, 2002, Cat. No. 0402 3 274 250

ISSN 0946 1310